Detecting P-waves in streaming seismic data using a hidden Markov model

J. Ray[1] , L. Burks[2] and C. Young[3],

[1]Sandia National Laboratories, Livermore, CA, [2]One Concern, Inc, Palo Alto, CA and [3]Sandia National Laboratories, Albuquerque, NM

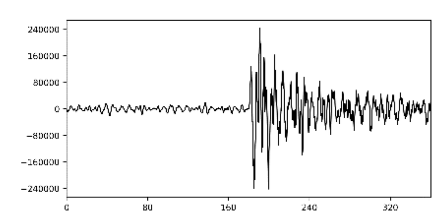
OBJECTIVE

Develop a method to fuse outputs of diverse seismic wave arrival detectors (algorithms) to improve performance i.e., reduce false positive and negative rates

- Use hidden Markov models (HMM)s to implement the fusion scheme
- Test on P-wave arrivals recorded at WB2, Warramunga IMS stations, NT, Australia

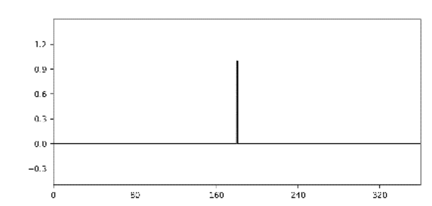
BACKGROUND

- A raw seismogram may show an abrupt commencement of activity (depending on SNR), indicating the arrival of a seismic wave



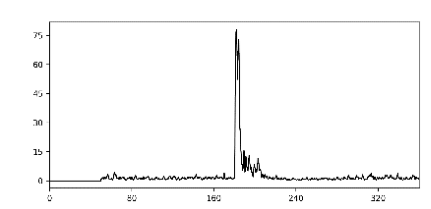
Waveform

- An analyst picks a time as the time of arrival of the seismic wave



Analyst's pick

- Detector algorithms like STA/LTA compute a function, derived from the seismic waveform, that has a "spike"
- The timing of the "spike" should be close to the arrival time



STA/LTA output

- The seismic waveform may have to be bandpass filtered first to enhance the signal relative to background noise
- Multiple detection algorithms add robustness: It is unlikely that a non-seismic (confounding process) will cause *all* detector algorithms to misfire at the same time
- Thus a simultaneous "spike" (a large change in detector function value) across multiple detector algorithm should indicate a seismic wave arrival with high confidence
- It should also reduce false positive (FP) and false negative (FN) rates
- But how to harness the *joint* predictive power of multiple detector algorithms?
- **Premise:** The fusion could be performed using hidden Markov models (HMMs)

FORMULATION

The Hidden Markov Model (HMM)

- Consider a system that occupies one of N states at time *t*, and moves to a new state at time *t+1*
- $H_t = \{h_i\}$, i = 1…N, are the probabilities that the system is in state *i* at time *t*
- Consider, too, that the system evolves as H_{t+1} = [P] H_t
  - Here [P] is a N x N matrix
  - p_{ij} is the probability of transitioning from state 'i' to state 'j' in one timestep, and is constant over time
  - This is a **Markov** system
- The state of the system is not seen (i.e., ***hidden***), but it causes observable phenomenon
- Let the observed state of the system at time t be one of K states
  - Let $O_t = \{o_k\}$, k = 1…K be the probabilities that the system is in observed state 'k'
  - Also O_t = [M] H_t, [M] is a K x N matrix
  - m_{ki} is the probability of hidden state i causing observed state k
- Such a discrete-time, discrete-state model for a time-dependent system is called a hidden Markov model (HMM)

Learning HMMs

- Given a long sequence of true internal states, one can learn the *transition* matrix [P]
- Given a long sequence of observed states, one can learn the emission matrix [M] from H_t
- If only observed states are available, along with guesses [P*] and [M*], one can:
  - Adjust [P*] to get [P] and [M*] to get [M]; also infer internal states H_t

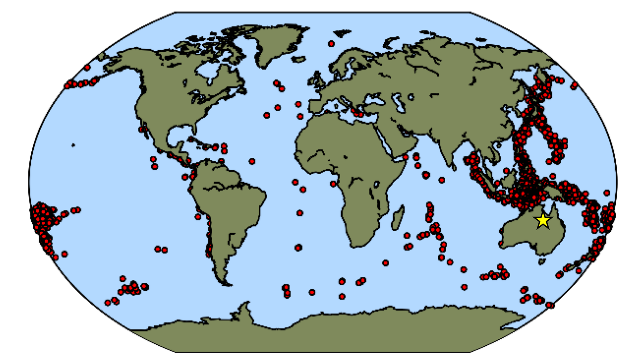
Fusion model

- The seismogram is supposed to be a binary system, capable of being in S = $\{s_1, s_2\}$, s_1 = noise or s_2 = seismic signal
- We have defined 10 observables. Each observable has many measured values which is represented as a Gaussian
- 3 observables are outputs of seismic wave detection algorithms. They should have a spike at the arrival time
- The other 7 observables also assume very different values for noise versus P-waves, e.g. rectilinearity. They help recognize the seismic signal.

FEATURIZING THE DATA

Data

- The dataset contains 835 P-wave arrivals as measured by the 3C station WB2 at Warramunga, NT, Australia
- The sources of these arrivals are events at regional and teleseismic distances from WB2; an analyst picked the arrivals manually

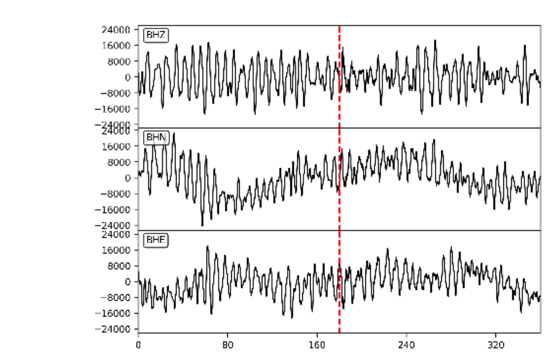


Map of WRA station (yellow star) and seismic events (red circles)

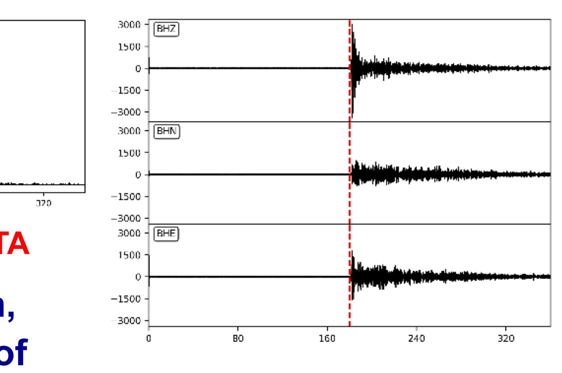
Features

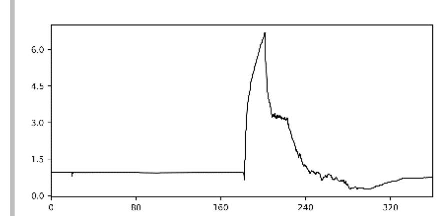
- The waveform was bandpass-filtered in (1.5, 3), (3, 6), (2, 5) & (6, 12) Hz bands
- In each band, the classic STA/LTA, recursive STA/LTA and Z-statistic algorithms provided the 3 observables with spikes to denote arrivals

Raw seismogram

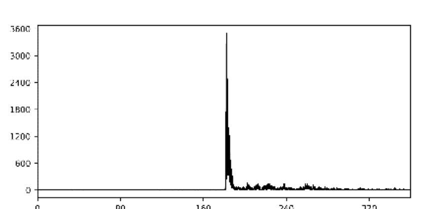


Seismogram bandpass filtered 2-5 Hz



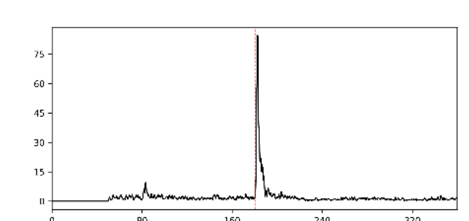


Z-statistic

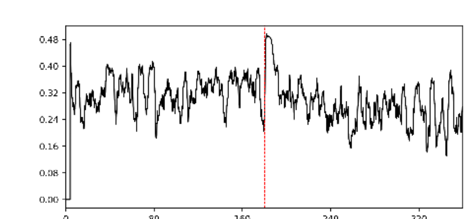


Classic STA/LTA

- In each band we computed polarization, rectilinearity, planarity, azimuth, angle of incidence and ratio of vertical to total power
  - These features help discriminate between noise and P-waves
  - The choice of features is dictated by their use in a streaming context – they are computationally inexpensive
- We also compute the area, in different frequency bands, under the PSD curve computed from the sonogram



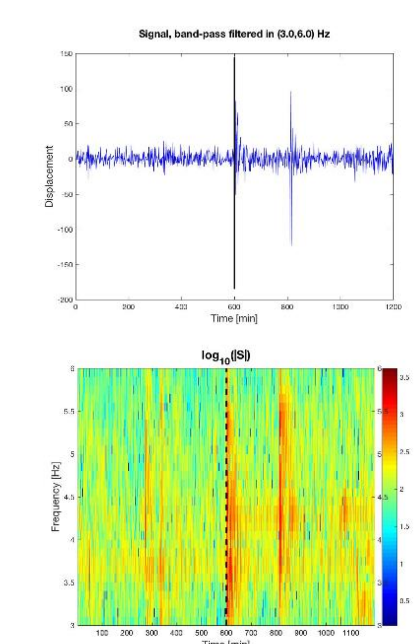
Polarization



Vertical to total power ratio

Constructing and testing HMMs

- The internal state of the HMM was computed by setting the state to 1 over a 4-second duration starting from the pick time.
- The HMM was trained over 80% of the dataset and tested over the remaining 20%
- The observed data of the test set was used to infer the internal (hidden) state (0-1 rise)
- Predicted arrival times were compared with the analyst time for the test waveforms



Waveform & spectrogram

RESULTS

- Two Gaussian HMMs were trained using 10-fold cross-validation
- One HMM was trained using the 3 detector algorithms' outputs as the observables
- The second also had the other observables (rectilinearity, etc.)
- If the HMM's predicted arrival time was within a few seconds of the analyst's, it was deemed a true detection

Detections		HMM (Detection algorithms only)		Classic STA/LTA		Z-Statistic	
		True	Missed	True	Missed	True	Missed
Expert-picked	True	774	61	753	82	219	616
	False	457		2879		1187	

- The performance of the HMM-based fused detector is summarized using the precision & recall
- TD: True Detection; FD: False Detection; FM: False Misses

- Precision = $\frac{TD}{TD+FD}$
- Recall = $\frac{TD}{TD+FM}$

	HMM (Detection algorithm only)	Classic STA/LTA	Z-Statistic
Precision	0.63	0.21	0.16
Recall	0.92	0.90	0.26

CONCLUSIONS

- The fusion of the 3 detector algorithms improved precision and recall compared to the individual algorithms
- Polarization etc., when used as features improved recall slightly, at the expense of recall
- However precision is still very low, and there are lots of FP.
- **Future work:** Train two HMMs, one each for noise and P-waves, and model-select between the two for a signal

Acknowledgements

The authors acknowledge the support of the National Nuclear Security Administration Office of Defense Nuclear Nonproliferation Research and Development for funding this work.

For additional information, please contact:

Jaideep Ray, Sandia National Laboratories, jairay@sandia.gov

SAND2018-13664C